



Toward New Indicators of a Journal's Manuscript Peer Review Process

Henk F. Moed*

Sapienza University of Rome, Rome, Italy

Journal impact factor is among the most frequently used bibliometric indicators in scientific-scholarly journal and research assessment. This paper addresses the question as to why this indicator has become so attractive and pervasive. It defends the position that the most effective way to reduce the role of citation-based journal metrics in journal and research assessment is developing indicators of the quality of journals' manuscript peer review process, based on an analysis of *this process itself*, as reflected in the written communication between authors, referees, and journal editors in electronic submission systems. This approach combines computational linguistic tools from the domain of "digital humanities" with "classical humanistic" text analysis and a profound knowledge of the manuscript peer review and the publication process.

Keywords: bibliometrics, indicator development, journal impact factor, manuscript peer review, research assessment

INTRODUCTION

Securing a political basis for academic research is a principal concern and a joint responsibility of (supra) national political domain and the academic research community. National research assessment exercises, performance-based funding, and assessments within institutions are important elements in this process. They increasingly need validated information and valid assessment methodologies in which indicators play a key role.

Despite the critique in the San Francisco Declaration on Research Assessment (DORA) published about 5 years ago, journal impact factors (JIFs) are still heavily used by research managers, evaluators, and publishers in the assessment of individual researchers and scientific journals. The DORA Manifesto (DORA, 2009) stated that the use of journal-based metrics must be eliminated in evaluation of individuals and greatly reduced in journal promotion (Van Noorden, 2013).

The JIF, a citation-per-article indicator, calculated at a level of a scientific journal, was developed by Eugene Garfield with the purpose of assessing a journal's information utility and correcting for the size of its annual volume and was used as a tool to monitor the coverage of his Science Citation Index (Garfield, 1972). Nowadays, they are not only used in librarians' or researchers' assessment of journals but also in journal editorial management and in setting targets in contracts between publishers and journal editors. They are also used in the assessment of the performance of individual researchers and groups. Its validity and utility has been discussed in numerous publications (e.g., Adler et al., 2008; Vanclay, 2012).

Although usage information on bibliometric indicators is not comprehensively collected and analyzed, it seems appropriate to assume that the JIF is among the most frequently used indicators in a journal or research assessment context. In Section "The Attractiveness of Journal Impact Factor," the question is addressed why this indicator has become so popular and pervasive. Four aspects are considered in an assessment of its attractiveness: visibility, availability, conceptual simplicity, and

OPEN ACCESS

Edited by:

Stasa Milojević,
Indiana University Maurer School of
Law, USA

Reviewed by:

Stefanie Haustein,
Université de Montréal, Canada
Kevin Boyack,
SciTech Strategies Inc., USA

*Correspondence:

Henk F. Moed
hf.moed@gmail.com

Received: 05 February 2016

Accepted: 20 July 2016

Published: 02 August 2016

Citation:

Moed HF (2016) Toward New
Indicators of a Journal's Manuscript
Peer Review Process.
Front. Res. Metr. Anal. 1:5.
doi: 10.3389/fрма.2016.00005

utility in terms of whether they provide information, which users find relevant.

Taking into account attractive features of this indicator highlighted in Section “The Attractiveness of Journal Impact Factor,” Section “A New Approach” proposes the application-driven development of a new research assessment tool, namely, indicators of the quality of a journal’s manuscript peer review process. Such a tool is not yet developed, and no blueprint is currently available. Its development requires a substantial amount of R&D in a long-term project. But the author of this paper would like to bring this approach to the attention of the scholarly community and defend the need to further develop it. The paper sketches the main lines of its development and shows how it embodies attractive features. Finally, Section “Concluding Remarks” highlights key differences in the landscape of scientific information providers and users between the 1950s and recent years and draws implications for the role of metrics developers.

During the past 50 years huge changes took place when the research process entered the modern, computerized, or digital age. The computerization of the research process does not merely relate to the collection of research data and the development of research methods but also to scientific information retrieval, scholarly communication and publishing, collaboration, and interaction between researchers and the wider public (Nielsen, 2011). The new approach aims to fully exploit the potential of this process, especially the use of computational linguistic approaches, and of advanced online information retrieval tools.

It must be underlined that the quality of the manuscript peer review process is definitely not the only relevant aspect of journal performance and that an approach directed toward the development of more direct indicators of the quality of this process does *not* aim to ban all citation-based indicators from the domain of journal performance assessment.

THE ATTRACTIVENESS OF JOURNAL IMPACT FACTOR

The Thomson Reuters ranking of journals by impact factor and subject category has become a standard not only in journal evaluation but also in research assessment of individuals, groups, and departments. Why did this happen? A series of factors with a possible influence may be identified. First, JIFs have always been very *visible*. Eugene Garfield published an essay on journal assessment in the journal *Science* (Garfield, 1972). They were flags of the Institute for Scientific Information. Second, JIF values are easily *available*. The lists of journals by subject category ranked by impact factor included in the Institute for Scientific Information’s (ISI, currently Thomson Reuters) journal citation reports (JCR) have always had a wide circulation.

Third, its definition is from a *conceptual* point of view, at least at first sight, relatively simple. It is an average, a type of statistic most users are familiar with. A discussion of technical statistical issues, such as the skewness of citation distributions (e.g., Seglen, 1994), the problem of “free” citations (Moed and Van Leeuwen, 1996), subject field biases (e.g., Vanclay, 2012), and of alternative journal impact measures (Pinski and Narin, 1976; West et al., 2008;

González-Pereira et al., 2010; Moed, 2010; Waltman et al., 2013), falls outside the scope of this discussion paper.

Finally, researchers and librarians find the journal information in the JCR *useful*, as the JCR was the first information product giving users a unique, integral picture of thousands of journals covering all major disciplines. In the assessment of research performance of individuals, groups, and institutions, JIFs are considered useful because users assume that they give an indication of the quality of a journal’s peer review process. Their base assumptions are, first, that publishing in journals with a rigorous manuscript peer review process is a valid quality marker and, second, that the best available indicator of the quality of this process is a measure based on citation impact.

Proponents of this type of use do *not* necessarily claim that JIFs are good “predictors” of the citation impact of individual articles, and many may agree with Garfield (1996), Seglen (1994), and Adler et al. (2008) who strongly criticize this claim. What they do believe is that large differences exist in the quality of manuscript peer review processes among journal editorial boards and that it is, from a policy viewpoint, appropriate to reward those authors who expose their manuscripts to critical referees applying high, internationally accepted quality standards.

But, what is the empirical basis of this claim that the JIF is a good indicator of the quality of the manuscript peer review process? Since direct indicators of the quality of peer review process are unavailable, empirical tests tend to be based on indirect, only partially valid indicators or proxies, such as manuscript rejection rates, as was done by Sugimoto et al. (2013). Also, studies in the past have validated citation-based journal indicators with perceptions of experts of peers on particular sets of papers or on journals as a whole. But, the correlation between these two measures tends to be weak. For instance, Bornmann and Leydesdorff (2012) obtained, in a set of 125 articles, a weak rank correlation coefficient of 0.3 between the impact factor of the journal in which they were published and the F1000 Article Factor, a measure based on post-publication article ratings given *via* a social media platform by several thousands of senior scientists in biology and medicine.

These proxies can be assumed to be to some extent positively related with the quality of the manuscript peer review process, although this is more obvious for journal manuscript rejection rates than it is for post-publication peer review. But, all suffer from severe biases as well. Journal citation rates are influenced by differences in citation practices among subject fields, by a journal’s visibility, availability, and prestige; moreover, impact and quality do not necessarily coincide (e.g., Moed, 2005). Manuscript rejection rates may be affected by author self-selection. Researchers’ perceptions may be influenced by a journal’s reputation and a post-publication review, by the time delay between post- and pre-publication review, and by its actual impact since its publication. If the quality of a journal’s manuscript peer review process is considered so important in research assessment, why not make an attempt to develop more direct indicators of this aspect? In the next section, a start will be made.

It is important to note that this does not mean that journal quality is or should be the sole aspect to be assessed in a research assessment. Other aspects, including “actual” citation impact or, at the non-bibliometric side, for instance, contribution to

innovation may be at least as important or even more relevant. The author of this paper wishes to take a neutral position as regards the role journal quality should play in a research assessment exercise. After all, the design of an assessment very much depends upon the assessment context. Indicators suitable in one context may be inappropriate in another. Based on the notion of a multidimensional research assessment matrix (AUBR, 2010; Moed and Halevi, 2015), the choice of an assessment methodology depends upon a series of factors: What is the unit of assessment? What is the aspect to be assessed? What is the objective of the process? And what is the state of the system (the total set of units of assessment) to which the assessment applies?

More direct indicators of the manuscript peer review process are urgently needed also because, nowadays, JIFs seem to have pervaded the entire scientific publication process. They are not only used in librarians' or researchers' assessment of journals but also in journal editorial management and in setting targets in contracts between publishers and journal editors-in-chief. Its dependence of the scientific publication process upon impact factors is so strong, that this process cannot do without these any more, and actually, in a positive feedback loop, further increases their importance. This is perhaps a main explaining factor of their pervasiveness.

A NEW APPROACH

I defend the position that the most effective way to reduce the role of citation-based journal metrics in journal and research assessment is the development of indicators of the quality of journals' manuscript peer review process, based on an analysis of *this process itself*, rather than on *proxies*, such as citation-based measures or manuscript rejection rates. To the extent that research evaluation agencies consider journal quality and especially the quality of its review process a relevant criterion in the assessment of individuals or groups and are interested as to whether a researcher under assessment has submitted his or her articles to a journal with a serious referee procedure and well instructed reviewers, these agencies would profit from more direct measures of journal quality, and indicator developers should make an attempt to develop these.

In a research project aimed to develop indicators of the journal manuscript review process, computational linguistic tools from the domain of "digital humanities" are useful, but "classical humanistic" text analysis and a profound knowledge of the manuscript peer review and the publication process are essential as well. The project would have at least *two* phases. A *first* phase involves the development of a conceptual *model* of manuscript peer review, including the construction of referee report profiles and communication modes between referees, authors, and editors. In this phase, a conceptual analysis is conducted of a sample of actual referee reports for a number of sources from different subject fields.

In the *second* phase of the project, when at least a first version of the model developed in the first phase is available, data mining is carried out of large numbers of electronic submissions. A linguistic analysis is conducted of peer review reports and the communication between authors, referees, and editors using natural language

processing and other computational linguistic techniques. It is in this phase that a statistical analysis of large datasets explores the construction of indicators of the peer review process, not only at the level of individual submissions but also at the level of journals and subject fields. The outcomes of such exploration may lead to adjustments in the model developed in the first phase.

In the current stage, concrete examples may narrow the perspective, or simply create misunderstandings, and violate the openness that is needed during the start-up phase. On the other hand, not giving any examples would make this discussion paper less convincing. Therefore, I give a possible line of inquiry. But, first, I wish to underline that there are many actors in manuscript peer review: authors, reviewers, journal editors, editorial boards, and journal publishers. The ultimate goal of the research project I propose is *not* to evaluate or rank particular actors according to specific quantitative measures but rather to improve the quality, efficiency, and effectiveness of the manuscript peer review process itself.

Two relevant issues would be as follows: has the reviewer read the manuscript sufficiently thoroughly? And: which assessment criterion has the reviewer actually applied? I start with the second issue. This inquiry is not only interested in general criteria such as "methodological soundness" but especially in the way in which these are operationalized in a reviewer's written report. To have more insight, evaluative statements should be identified in the text as well as the "standards" they apply, either explicitly or implicitly. If one would do so for a number of reports for a particular journal, one could examine the degree of agreement among reviewers as regards the standards they apply, and if one collects data for more journals, make comparisons across journals. Instructions or recommendations by journal editors to their reviewers should be taken into account as well. A second analysis would focus on the amount or degree of detail of the information provided by the reviewer about the manuscript under review. One could formulate and further validate the hypothesis that review reports that apply vague standards or fail to apply assumed key standards, and that contain no reference to the text of the manuscript under review, are of less quality than those adopting a series of clear assessment criteria and supporting their assessment by citing text passages or tables from the manuscript.

This proposed research project has to maintain strict *anonymity* with respect to the analyzed manuscripts and referee reports. For instance, argumentation structures need to be formalized and detached from the concrete information on the cases from which they are derived. This approach can be applied at a large-scale only if, especially in the second phase of the project, journal publishers are prepared to participate by disclosing under strict conditions of confidentiality parts of their online submission systems. But, exploratory studies aimed to develop a base methodology, and to show the feasibility of the approach, could focus on publication sources that tend to be much more easily available, namely, on peer reviewed *proceedings of international conferences*.

It must be noted that the quality of the manuscript peer review processes is a research topic in its own right, regardless of whether it aims to contribute to the development of better indicators of journal quality. The outcomes of the proposed study could further enhance the transparency of the manuscript referee process, also

for submitting authors. Most importantly, in my view, they could help educating and training new reviewers. Next, they could help assessing the effect of peer review upon manuscript quality. This could provide information that can be used to demonstrate the added value of the process. Finally, it could contribute to further operationalizing the multidimensional concept of quality of a journal's peer review process and develop indicators that can be used to monitor and further improve this process, jointly with advanced online tools editors and reviewers need in their tasks, such as, for instance, plagiarism detection tools. In this way, these indicators could potentially be used as an alternative of the JIF as a more direct measurement of the quality of a journal's manuscript peer review process.

CONCLUDING REMARKS

Beyond any doubt, indicators applied in assessment process must have a sufficiently high level of accuracy, validity, and methodological sophistication. In this respect, much progress has been made during the past decennia. But, in the type of tool I propose, the trade-off between methodological sophistication and usability for large user groups should be in favor of the latter. Sophisticated indicators are particularly useful as *research tools* in testing specific hypotheses in quantitative science and technology studies, but are not necessarily useful assessment tools for a wide user group. Sophisticated indicators can be used to validate

simplified indicator variants derived from them, which are more easily intelligible and useful for large groups of users.

Since Eugene Garfield introduced the JIF as an “objective” tool in a journal coverage policy of his citation index independently of journal publishers, the landscape of scientific information providers and users has changed significantly. While, on the one hand, politicians and research managers at various institutional levels need valid and reliable fit-for-purpose metrics in the assessment of publicly funded research, there is, on the other hand, a tendency that metrics increasingly become a tool in the business strategy of companies with product portfolios, which may include underlying databases, social networking sites, or even metrics products. This may be true both for “classical” bibliometric indicators and for alternative metrics. In metrics development, competent developers must apply rigorous scientific criteria in the examination of the validity, reliability, and utility of new indicators.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and approved it for publication.

ACKNOWLEDGMENTS

The author wishes to thank two anonymous referees for their useful comments on two earlier versions of this paper.

REFERENCES

- Adler, R., Ewing, J., and Taylor, P. (2008). *Citation Statistics. A report from the International Mathematical Union*. Available at: www.mathunion.org/publications/report/citationstatistics0
- AUBR. (2010). *Expert Group on the Assessment of University-Based Research. Assessing Europe's University-Based Research*. Available at: https://ec.europa.eu/research/science-society/document_library/pdf_06/assessing-europe-university-based-research_en.pdf
- Bornmann, L., and Leydesdorff, L. (2012). *The Validation of (Advanced) Bibliometric Indicators through Peer Assessments: A Comparative Study Using Data from InCites and F1000*. Available at: <https://arxiv.org/pdf/1211.1154>
- DORA. (2009). *San Francisco Declaration on Research Assessment*. Available at: <http://www.ascb.org/dora/>
- Garfield, E. (1972). Citation analysis as a tool in journal evaluation. *Science* 178, 471–479. doi:10.1126/science.178.4060.471
- Garfield, E. (1996). How can impact factors be improved? *Br. Med. J.* 313, 411–413. doi:10.1136/bmj.313.7054.411
- González-Pereira, B., Guerrero-Bote, V. P., and Moya-Anegón, F. (2010). A new approach to the metric of journals' scientific prestige: the SJR indicator. *J. Informetrics* 4, 379–391. doi:10.1016/j.joi.2010.03.002
- Moed, H. F. (2005). *Citation Analysis in Research Evaluation*. Dordrecht: Springer.
- Moed, H. F. (2010). Measuring contextual citation impact of scientific journals. *J. Informetrics* 4, 265–277. doi:10.1016/j.joi.2010.01.002
- Moed, H. F., and Halevi, G. (2015). Multidimensional assessment of scholarly research impact. *J. Am. Soc. Inf. Sci. Technol.* 66, 1988–2002. doi:10.1002/asi.23314
- Moed, H. F., and Van Leeuwen, T. N. (1996). Impact factors can mislead. *Nature* 381, 186. doi:10.1038/381186a0
- Nielsen, M. (2011). *Re-inventing Discovery: The New Era of Networked Science*. Princeton, NJ: Princeton University Press.
- Pinski, G., and Narin, F. (1976). Citation influence for journal aggregates of scientific publications: theory, with application to the literature of physics. *Inf. Process. Manag.* 12, 297–312. doi:10.1016/0306-4573(76)90048-0
- Seglen, P. O. (1994). Causal relationship between article citedness and journal impact. *J. Am. Soc. Inf. Sci.* 45, 1–11. doi:10.1002/(SICI)1097-4571(199401)45:1<1::AID-ASI>3.0.CO;2-Y
- Sugimoto, C. R., Larivière, V., Ni, C., and Cronin, B. (2013). Journal acceptance rates: a cross-disciplinary analysis of variability and relationships with journal measures. *J. Informetrics* 7, 897–906. doi:10.1016/j.joi.2013.08.007
- Van Noorden, R. (2013). *Scientists Join Journal Editors to Fight Impact-Factor Abuse*. *Nature News Blog*. 16 May 2013. Available at: <http://blogs.nature.com/news/2013/05/scientists-join-journal-editors-to-fight-impact-factor-abuse.html>
- Vanclay, J. K. (2012). Impact factor: outdated artefact or stepping-stone to journal certification? *Scientometrics* 92, 211–238. doi:10.1007/s11192-011-0561-0
- Waltman, L., Van Eck, N. J., Van Leeuwen, T. N., and Visser, M. S. (2013). Some modifications to the SNIP journal impact indicator. *J. Informetrics* 7, 272–285. doi:10.1016/j.joi.2012.11.011
- West, J., Althouse, B., Rosvall, M., Bergstrom, T., and Bergstrom, C. (2008). *Eigenfactor: Detailed Methods*. Available at: www.eigenfactor.org/methods.pdf

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Moed. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.